CIRANO
Allier savoir et décision

2016s-47

Learning in a Bandit Game and Technology Choice

Kaywana Raeburn, Jim Engle-Warnick, Sonia Laszlo, Jian Li

Série Scientifique/Scientific Series

2016s-47

Learning in a Bandit Game and Technology Choice

Kaywana Raeburn, Jim Engle-Warnick, Sonia Laszlo, Jian Li

Série Scientifique
Scientific Series

Montréal
Août/August 2016

CIRANO
Allier savoir et décision

Centre interuniversitaire de recherche en analyse des organisations

CIRANO

Le CIRANO est un organisme sans but lucratif constitué en vertu de la Loi des compagnies du Québec. Le financement de son infrastructure et de ses activités de recherche provient des cotisations de ses organisations-membres, d'une subvention d'infrastructure du ministère de l'Économie, de l'Innovation et des Exportations, de même que des subventions et mandats obtenus par ses équipes de recherche.

CIRANO is a private non-profit organization incorporated under the Quebec Companies Act. Its infrastructure and research activities are funded through fees paid by member organizations, an infrastructure grant from the ministère de l'Économie, de l'Innovation et des Exportations, and grants and research mandates obtained by its research teams.

Les partenaires du CIRANO

Partenaires corporatifs
Autorité des marchés financiers
Banque de développement du Canada
Banque du Canada
Banque Laurentienne du Canada
Banque Nationale du Canada
Bell Canada
BMO Groupe financier
Caisse de dépôt et placement du Québec
Fédération des caisses Desjardins du Québec
Gaz Métro
Hydro-Québec
Industrie Canada
Intact
Investissements PSP
Ministère de l'Économie, de l'Innovation et des Exportations
Ministère des Finances du Québec
Power Corporation du Canada
Rio Tinto
Ville de Montréal

Partenaires universitaires
École Polytechnique de Montréal
École de technologie supérieure (ÉTS)
HEC Montréal
Institut national de la recherche scientifique (INRS)
McGill University
Université Concordia
Université de Montréal
Université de Sherbrooke
Université du Québec
Université du Québec à Montréal
Université Laval

Le CIRANO collabore avec de nombreux centres et chaires de recherche universitaires dont on peut consulter la liste sur son site web.

ISSN 2292-0838 (en ligne)

Learning in a Bandit Game and Technology Choice*

Kaywana Raeburn†, *Jim Engle-Warnick*‡, *Sonia Laszlo*§, *Jian Li***

Abstract

We present a decision-making experiment, conducted in the field, that explores the extent to which learning in a bandit game predicts technology choice on the farm. We find evidence of heterogeneity of learning in the bandit game, including overweighting of winning draws and Bayesian updating. Our results suggest that learners who overweight recent draws are more likely to have adopted new farm crops within the previous year.

Keywords: Bandit Games, Technology Choice, Field Experiment.

Codes JEL/JEL Codes: C90, O33, Q16.

---

* This work was carried out with the aid of a grant from the International Development Research Centre (IDRC), Ottawa, Canada.
† McGill University. Corresponding author: kaywana.raeburn@mail.mcgill.ca.
‡ McGill University and CIRANO.
§ McGill University and CIRANO.
** McGill University.

1 Introduction

In this paper we explore the connection between learning and technology choices in an artefactual field experiment conducted with farmers in Guyana. In our experiment, we introduce a multi-period one-armed bandit game, which we turn into an instrument to measure behavior. We use choice patterns in the bandit game to identify different styles of learning among the farmers. We then correlate estimated learning styles with farmer technology adoption and stopping choices.

In our framework, subjects play multiple three period games in which they make a series of choices between a known arm with certain amounts of earning and an unknown arm, which results in a win or loss of with unknown probabilities. Public revelation of the outcomes from the unknown arm can be used to update beliefs about the likelihood of success of the risky outcomes. We construct a number of indices to measure degrees of correctness of decision-making relative to proper Bayesian updating. We then use choices in the games to estimate and categorize individuals into learning styles.

We find heterogeneity in decision-making, including overweighting of winning draws, and a small but significant amount of Bayesian learning. We correlate these learning styles with survey data on farmers sociodemographics and on input choices on their farms. Our results suggest that individuals who overweight recent draws are more likely to have adopted new farm inputs within the previous year. This suggests that focusing on Bayesian learning in adoption models may lead to inaccurate predictions, and that it is important to consider potential differences in how individuals learn and the subsequent effects on decision-making.

Theories of learning are important in dynamic contexts, and many economic models of learning, including Bandit games, involve Bayesian updating. However, it is well-known that people often fail to update like Bayesians. For example, it has been found that people both underweight and overweight the likelihood (Grether, 1980). For another example, it has been found that people see negative serial correlation when it does not exist (the gambler's fallacy (Croson and Sundali, 1995), and positive serial correlation when it does not exist (the hot hand fallacy (Camerer, 1989). Thus it is not enough to simply test behavior against learning theories. It is also important to understand which in a set of possible alternative learning behaviors is being exhibited. The answer to this question has implications for theory as well as for policy. We introduce an experimental method to identify a learning type.

Theories of learning are often applied to specific contexts. In developing countries, for example, farmers tend to be slower to adopt new technologies. While many determinates of this tendency have been studied, including risk preferences (Binswanger, 1980; Knight, et al., 2003; Liu, 2013) and ambiguity preferences (Ross, et al., 2012; Engle-Warnick et al., 2011), the role of learning is also prevalent (Foster and Rosenzweig, 2010; Conley and Udry, 2010).[1] Farmers can learn through a learning-by-doing mechanism, in which individuals use information from their own experience to update beliefs about new technologies, or through social learning, where information from other individuals is used to update beliefs (Foster and Rosenzweig, 1995). Research that explores either of these frameworks tends to concentrate on Bayesian updating in the learning process. Our work is in the vein of Barham et al. (2015), who suggest that farmers in developed countries use learning rules that are very heterogeneous, with only a small proportion using Bayesian updating.

This paper provides empirical evidence on the real-life consequences of different learning styles. In it we uniquely turn a bandit game into a behavioral instrument to classify learning styles, and use the instrument to shed new light on an important and well-studied problem. Our results, which show that people who over-react to new information are most likely to try new inputs, suggest a potential new take on the adoption of new technologies.

[1]The learning-by-doing model of Jovanovic and Nyarko (1996), for example, is an alternative learning model specifically for technology choice.

2 The Model

2.1 Bandit Games

In a multi-armed bandit problem a decision-maker chooses between k possible arms or experiments. The choice of arm i results in an observation/reward from experiment i. The observation may provide useful information for a future choice between the arms. The goal is to maximize the present value of the discounted stream of rewards. The tradeoff is between exploration (finding the best arm) and exploitation (choosing the arm thought to have the best playoff). A simplification of the multi-armed bandit problem is a "one-armed" bandit problem with two arms in which one arm has a known i.i.d. distribution of returns.

The bandit framework is used extensively in studies of information acquisition and learning by economic agents. In economics, for example, it has been used to model market pricing and learning (Rothchild, 1974; Rustichini and Wolinsky, 1995), labor search and matching (Jovanovic, 1979), corporate finance and asset pricing (Bergemann and Hege, 1998; 2005) and technology adoption (Copeland, 2007).

Bandit games have been empirically tested in the experimental laboratory. Banks, et al. (1997) conduct laboratory experiments to analyze and compare decision-making in two simple bandit problems - a one-armed bandit problem and a two-armed bandit problem. They find that behavior in the two environments are qualitatively different and in the theoretically predicted direction. Gans, et al. (2007) investigate how well simple models of discrete choice match actual performance in bandit experiments. They find that simpler learning rules match best while the more complicated Bayesian learning perform worst. Boyce, et al. (2015) use a laboratory experiment to examine whether individual behavior is consistent with the Nash equilibrium predictions of a one-armed bandit game with information spillover. They find that while there is a bias towards under-experimentation, there is a significant strategic effect.

This paper builds on this literature by using a bandit game as an instrument to measure behavior in an artefactual field experiment (Harrison and List, 2004) with farmers in a developing country. In our framework, we use a three-period one-armed bandit problem to observe how people learn, and we examine whether those choices can be modeled by simple learning rules and whether those learning rules have predictive power with regard to actual choices on their farms.

2.2 The Model

Our model of learning is a "one-armed" bandit game in which a decision-maker (DM) selects between two alternatives: a known arm (K) that returns a constant amount of earnings, and an unknown arm (A) that has binary outcomes labelled success and failure with unknown probability.

2.2.1 Beliefs and Probabilities

Suppose the DM has a 50-50 prior, μ_0, that the likelihood of a success for selecting the unknown arm is either θ or $1-\theta$. If her beliefs are updated in a Bayesian manner, then after t draws the posterior, μ_t, conditional on observing k successes and $t-k$ failures is:

$$\mu_t = Pr(\theta|k\ successes, t-k\ failures) = \frac{0.5\theta^k(1-\theta)^{t-k}}{0.5\theta^k(1-\theta)^{t-k} + 0.5(1-\theta)^k\theta^{t-k}}$$

2.2.2 Value Functions

Consider the case of a three-period game. Let $t = 0, 1, 2$ indicate the period, S represent success and F represent failure of the unknown arm, μ_t be the belief at the beginning of the period, $p_t = \mu_t\theta + (1-\mu_t)(1-\theta)$ be the predictive probability of observing $x = S$ in period t, $u(x)$ be the per period utility function, and $x \in \{S, F\}$. For a given path of realizations of the unknown arm and fixed payoff for the known arm, λ, the value functions can be calculated as follows:

$$V_0 = max\{3u(\lambda),\ E_{\mu_0}[u(x) + V_{1,x})]\}$$
$$= max\{3u(\lambda),\ p_0\big(u(S) + V_1|x = S\big) + (1 - p_0)\big(u(F) + V_1|x = F\big)\}$$
$$V_1 = max\{2u(\lambda),\ E_{\mu_1}[u(x) + V_{2,x})]\}$$
$$= max\{2u(\lambda),\ p_1\big(u(S) + V_2|x = S\big) + (1 - p_1)\big(u(fF) + V_2|x = F\big)\}$$
$$V_2 = max\{u(\lambda),\ E_{\mu_2}u(x) = p_2u(S) + (1 - p_2)u(F)\}$$

2.2.3 Optimal Strategy

Let $s = (s_0, s_1, s_2)$ be a decision strategy where $s_t \ \epsilon \ \{K, A\}$ is the action taken at time t. Note that K represents selection of the known arm and A represents selection of the unknown arm. An optimal strategy exists which maximises the value function V_0. Strategy s^* is optimal if:

$V_0\{s^*\} > V_0\{s\}\ \forall s$ in the strategy space S.

The optimal strategy depends on the likelihood of success θ, the updated beliefs about the likelihood of success μ, and the earnings of the known arm λ.

3 Experimental Design

The experiment consists of a series of three-period one-armed bandit games as described above. In each game, participants choose between Option K, a fixed amount of earnings ($\$\lambda$) for sure, and Option A which consists of outcomes that are either success ($1500) or failure ($0) with unknown probabilities. In each game the actual likelihood of success of the unknown arm is either θ or $1 - \theta$.

Fixing the outcomes for success and for failure at 1500 and 0, our challenge was to jointly select values of θ (the likelihood of success) and λ (the value of the known arm) for a variety of optimal strategies. The experimental designed consisted of several different optimal strategies because of the possibility that the pattern of various strategies could simply be the result of a rule of thumb, rather than proper updating. Being mindful of the fact that values of θ closer to one induce faster learning, we spread the strategies the best we could across different values of θ.

Setting θ equal to four different values, $\{0.5, 0.7, 0.9, 1\}$, each representing increasing speed of learning with a draw from the unknown arm, for each value of θ we then selected several values of λ using simulations so that we obtained the following four optimal strategies.[2]

1. Choose the known arm in each period (KEP)
2. Choose the known arm after one failure observation ($K1F$)
3. Choose the known arm after two failure observations ($K2F$)
4. Choose the unknown arm in each period (AEP)

Using a range of values for λ, we evaluated V_0 for each strategy.[3] This resulted in each strategy being optimal for a specific range of values for the known arm λ conditional on the value of θ. The design is summarized in Table 1, where the third and the fourth columns show the range for λ for which a given strategy and value of θ are valid. The fourth column, labelled "Calibrated λ", presents the actual value used in the experimental design.[4] This resulted in the following optimal strategies for each θ.

Note that there are a total of ten parameter combinations in Table 1, represented by each row. Notice that for $\theta = 0.5$, where no updating can occur, we chose both strategies

[2]Bradt, et al. (1956) show that for one-armed bandit games with finite horizon and Bernoulli trials, if the known arm is optimal at any stage, then it is optimal to use that arm at all subsequent stages. Given this result, we focused on strategies for which there is a single switching point from the unknown arm to the known arm.

[3]We use the CRRA utility function $u(x) = \frac{x^{1-r}}{1-r}$ with coefficient $r = 0.5$.

[4]For $\theta = \{0.7, 0.9\}$, to calculate the fixed alternative for $K1F$ and $K2F$, we use the approximate midpoint of the ranges. The fixed alternative for KEP was approximated by adding the distance between the midpoint of the $K1F$ value and the upper bound to that upper bound. For $\theta = \{0.5, 1\}$ we used a (-275, +275) range around the switch λ of 375. The values are shown in column 4 in Table 1.

that select the same arm all three times to include in our experimental design. With the exception of choosing the unknown arm each time, which occurs once, every strategy occurs at least twice.

4 Experimental Procedures

4.1 Setting and Subject Pool

We held experimental sessions with farmers in communities in two administrative regions in Guyana. Ten sessions were held in East Berbice-Corentyne (Region 6) and seven sessions were held in Essequibo Islands-West Demerara (Region 3). All communities are rural villages where agriculture is one of the main livelihoods. The communities are quite heterogeneous in the agricultural goods produced and do not specialize in any particular crop, however cash crops such as fruits and vegetables are most commonly planted.

Recruitment of participants and organization of the locale for the experiment were conducted by our local field staff, which consisted of regionally-based employees of our local partner, the National Agriculture Research and Extension Institute (NAREI). All sessions were held in locations central to the surrounding communities. Each community was visited several days in advance to recruit subjects. Subjects had to be farmers of legal age (18 or above), reside in one of the adjacent rural communities and have basic literacy and numeracy skills. If a recruited subject was unable to attend on the day indicated, replacement participants were selected from a reserve list. In total, 295 subjects participated in the sessions, with session size ranging from 11 to 24 participants in Region 3 and 14 to 20 participants in Region 6.[5]

[5] Two participants had to leave the session prior to completing the experiment, two participants left before completing the exit survey and one participant did not participate in the incentivized part of the session due to religious reasons.

4.2 Session Details

We ran our sessions as laboratory experiments in the field, and experimental procedures were identical in both regions. Advised consent was obtained and instructions were read from a script by a single experimenter. Each session consisted of nine trials with three decision-making periods in each trial. Each trial corresponded to a specific value of θ, and in each period individuals made choice between the fixed earnings $\$\lambda$ and the unknown arm (Option A). In each trial, the distribution of θ was fixed for all periods. For $\theta = \{0.7, 0.9\}$, we had three trials of choices; for $\theta = \{1\}$, we had two trials of choices; and for $\theta = \{0.5\}$ we had one trial of choices.

At the beginning of the first period, we established the 50-50 prior that the likelihood of success was either θ or $1-\theta$ and the subjects were asked to choose between various $\$\lambda$'s and Option A. An outcome realization from Option A was then shown. Subjects made another choice between the $\$\lambda$'s and Option A in the second period,and then a second outcome realization from Option A was shown. Subjects made a final choice between the $\$\lambda$'s and Option A in the third period, then a third outcome realization from Option A was shown. The outcome realizations were drawn with replacement.

Before each trial, participants were given the two-page instrument (Figure 1) on which they recorded their decisions for that trial. On the first page, we presented an illustration of the unknown arm. There were two bags, each containing ten chips with values of either GUY\$1500 or GUY\$0. One bag contained a number of GUY\$1500 chips corresponding to the value of θ for the round, while the other bag contained a number of GUY\$1500 chips corresponding to $1-\theta$. For instance, the diagram in Figure 1 is an example from a trial in which $\theta = 0.7$. In the figure, one bag contains seven GUY\$1500 chips and three \$0 chips while the other bag contains three GUY\$1500 chips and seven GUY\$0 chips. To establish the 50-50 prior, we indicated above each bag that there was a 50% chance that the chips they will see come from that specific bag, and reinforce this in the instructions. The sequence of outcomes for each round was randomly chosen prior to the laboratory sessions for ease of implementation and was the same in all sessions.[6]

Subjects made their decisions in the experiment on the second page of the instrument. Each decision was a choice between the known payoff (λ) and the unknown arm, Option A. They played several games at once, each game corresponding to a specific value of λ. The decisions were made sequentially by rows. After the experimenter read instructions explaining the distribution of the bags in Option A, the participants made the decisions in the first row. An outcome from Option A was then shown to participants, following which they made the decisions in the second row of the instrument. Another outcome from Option A was shown and participants made the decisions in the third row of the instrument. Each row corresponds to a period in the trial.

A final outcome from Option A was then shown to the participants, who then made the decision on the first page of the instrument about which bag/distribution they believe the outcomes shown are from. This was repeated for each trial on a separate instrument. The trials were differentiated by the value of θ, and by values of the observations from the unknown arm. The value of θ and the outcomes from Option A are shown in Table 2. In total, participants made 72 decisions.[7]

After all trials were completed, a short exit survey was privately administered to each participant by our field staff. The survey collected information on demographics, farming practices and decisions and climate change perceptions and adaptation. We also measured risk and ambiguity attitudes using hypothetical (non-incentivized) choice instruments adapted from Engle-Warnick et al. (2009 and 2011).[8] Upon completion of the survey, subjects proceeded individually and in private to sit with the experimenter to determine their session earnings. The experiment lasted approximately one hour and the entire session including

[6] With the exception of one session in which Trial 5 was omitted due to time and light constraints.

[7] Note that all subjects saw the draw from the unkown arm whether they chose it or not. We did this to speed the learning process.

[8] The risk preference instrument requires subjects to choose among several binary pairs of gambles that vary in their level of risk. The more they choose from the relatively risky gamble, the more risk averse they are assumed to be. Similarly, the ambiguity instrument involves a series of decisions between several binary pairs of a risky but costly gamble and and ambiguous but costless gamble. The more often subjects chose to pay to avoid the ambiguous gamble, the more ambiguity averse they are assumed to be.

the survey and payment lasted two hours on average.

Subjects were given a show-up fee of GUY$1000 upon arrival to cover their transportation and opportunity cost for participation. We paid the show-up fee immediately to facilitate trust in the incentivized part of the experiment. Subjects were also paid for one of the decisions they made during the experiment. Each decision was numbered from 1 to 72. To determine which decisions they were paid for, each individual selected a chip from a bag containing chips numbered 1 to 72. For that decision, if they had chosen the known payoff λ, they were paid that amount. If they had chosen Option A, they were paid the amount of the chip from Option A that was revealed following that decision. If the decision was in row one, they were paid the value of the first chip shown; if the decision was in row two, they were paid the value of the second chip shown; and if the decision was in row three, they were paid the value of the third chip shown. Since there was the possibility of earning GUY$0 for the experiment result, we also paid an additional amount of GUY$500 at the end of the session for completing the survey. Subjects earned an average of GUY$700 (ranged from GUY$0 to GUY$1500) for the incentivized part of the experiment. Total earnings averaged GUY$2200.

5 Results

5.1 Exit Survey Descriptive Statistics

Table 3 presents the descriptive statistics of the farmer and household characteristics. The average participant in the experiment was about 44 years old, most likely to be of East Indian descent, married and more likely to be male (69%). The highest level of education received was primary school (46%), followed by secondary school (45%) and the remainder split between post-secondary schooling (6%) and less than primary (2%). Participants came from households with an average size of 4 individuals. While two-thirds of participants had flush toilet facilities only 37 percent used indoor taps as their source of drinking water. As an indicator of wealth, we also created an ownership index (ranging from 0 to 1) of five vehicle assets. The average of this index is 0.27.

Participants had an average of 20.5 years experience in farming and had farms which measure an average of 9.37 acres. Seventy-four percent of respondents owned their farm land. To get a measure of access to credit, we asked the farmers how easily they could obtain a loan or credit to make improvements on their farm on a scale of 1 (never) to 5 (always) from formal and informal sources. The median response was 3 for both sources, and the means were similar at 2.71 and 2.63 respectively. This suggests that on average farmers felt they had somewhat limited access to credit. Three-quarter of respondents received technical assistance from at least one source, while only 21 percent were member of a farming organisation. As expected the main crops planted by the farmers in our sample were fruits and vegetables. Table 4 provides a breakdown of the main crops reported to be cultivated. The five most common crops were pepper, bora (yardlong beans), okra, rice and boulanger (eggplant), four of which are cash crops with short growing cycles. Sixty-one percent of participants also raised livestock and 43 percent did work other that farming.

Figures 2 and 3 present histograms for the hypothetical risk and ambiguity preference measures elicited in the exit survey. The risk preference measure is the number of times the subject chose the safer of two gambles in four decisions and the ambiguity preference measure is the number of times the subject chose to pay to avoid the ambiguous gamble (selecting a costly known gamble with the same payoffs) in five decisions.[9] The risk measure is thus increasing in risk aversion and, similarly, the ambiguity preference measure is increasing in ambiguity aversion. For both measures there is heterogeneity in responses. The measures are also positively and significantly correlated (r = 0.4805) so that relatively more risk averse subjects also tend to be relatively more ambiguity averse.

[9] See Engle-Warnick, et al. (2011) for details on how these measures are constructed.

5.2 Experiment Results

5.2.1 Choices vs. Optimal Strategies

Each bandit game has a strategy which maximises the value of the game (V_0). We first construct an index which captures the degree of correct of decision-making by comparing each decision in each game to the corresponding decision in the optimal strategy. We compute a score ranging from 0 to 3 for each game played. For instance, for $\theta = 1.0$ and $\lambda = \$650$, the optimal strategy for the game is to choose λ in each period. Thus if the participant chose Option A in each period the score is 0; if he chose λ in one period only, the score is 1; if he chose λ in two periods,the score is 2; and if he chose λ in all periods the score is 3. With three periods and two possible options, there are eight potential sequence of choices that participants can make. One of these sequences corresponds to the optimal strategy and a score of 3, three of them would result in a score of 2, three would result in a score of 2, and one would result in a score of 0.

Figures 4, 5 and 6 show the distribution of the score (with the scores of 2 and 3 normalized by dividing by their number of occurances). From the figures, we see heterogeneity both within and between games. Most games with $\lambda = GUY\$375$ have scores closest to uniformly distributed, while in most games with LOW and HIGH λ, there is a mode at 3 correct decisions suggesting that there is increased difficulty in deciphering the MEDIUM λ games.

To explore this further we construct a weighted game-level index of correct decision-making using the following formula:

$$I_g = \frac{3 * (No.of3s) + 2 * \left(\frac{No.of2s}{3}\right) + 1 * \left(\frac{No.of1s}{3}\right) + 0 * (No.of0s)}{No.ofParticipants} \tag{1}$$

Table 5 shows the index for each game in the experiment by θ and λ. The mean game index is 0.965 which is significantly higher than 0.75 which is the index that would correspond to random game playing (uniform distribution of scores). As suggested by the table, there is heterogeneity between games, with the index ranging from 0.705 in Trial 7 (λ=HIGH, θ=0.9) to 1.306 (λ=LOW, θ=0.5) in Trial 9. Trials 2 and 6, which correspond to observations of 0, 0, 1500, have the lowest indices for the LOW λ; while rounds 4 and 7, which correspond to observations of 1500, 1500, 1500 have the lowest indices for the HIGH λ. This suggests that when the outside option is low, successive failures tend to lead to incorrect decision-making, while when the outside option is high successive successes tend to lead to incorrect decision-making. Thus both the value of the outside option and the luck of the draw appear to matter for correct decision-making.

Similar to the game index computed above, we construct an individual-level index of correctness as follows:

$$I_i = \frac{3 * (No.of3s) + 2 * \left(\frac{No.of2s}{3}\right) + 1 * \left(\frac{No.of1s}{3}\right) + 0 * (No.of0s)}{No.ofGames} \tag{2}$$

Figure 7 shows the distribution of this index. Again there is heterogeneity of behavior, with a mean of 0.964, which is significantly higher than the 0.75 of random game play.

5.2.2 Choices vs. Optimal Conditional on History

An alternative way to consider the qualiity of learning in the game is to consider each decision as autonomous and compare this with the optimal decision given the history of outcomes. Assuming a CRRA utility function with parameter $r = 1/2$ and Bayesian updating of posterior beliefs, we compute the optimal decision for each choice by comparing the expected utility of Option A with the expected utility of the fixed earnings λ. We then calculate the an index that captures the percentage of times the participants chose option A when they should have chosen option A and an index that captures the percentage of time participants chose the fixed earnings when they should have chosen the fixed earnings. These two indices together will capture the level at which participants are Bayesian in their updating. The indices are computed as follows:

$$Index_{A|A\ is\ Optimal} = \frac{No.\ of\ Times\ Chose\ A|A\ is\ Optimal}{No.\ of\ Times\ A\ is\ Optimal} \tag{3}$$

$$Index_{\lambda|\lambda\ is\ optimal} = \frac{No.\ of\ Times\ Chose\lambda|\lambda\ is\ Optimal}{No.\ of\ Times\lambda\ is\ Optimal} \tag{4}$$

Table 6 shows the means of these indices and the difference between the two. Both values are substantially and significantly less than 1, indicating that individuals are far from Bayesian in their their updating. The mean index of choosing Option A correctly is significantly greater that that of choosing λ correctly, suggesting that on average individuals made more correct decisions when they should 'experiment' versus when they should play 'safe'.

5.2.3 Learning Types

We next analyze learning at the individual level by fitting individual decisions to different learning styles (Gans, et al., 2007; Barham et al., 2014). While we have a small number (three) of observatons in each trial we have a total of 72 decisions for each participant.

We focus on the following five learning rules: Bayesian, first-1 (emphasis on the first draw), last-1 (emphasis on the previous draw), random, and status-quo (same decision each time). We utilise a logit framework to analyze the probability of choosing Option A in period t of game g defined by

$$Pr(Choose\ Option\ A) = \frac{e^{\beta x_{gt}}}{1 + e^{\beta x_{gt}}}, \tag{5}$$

where x_{gt} is the expected utility difference from choosing Option A in period t of game g and

$$x_{gt} = U(Win)p_{gt} + U(Lose)(1 - p_{gt}) - U(\lambda_g),$$

where $p_{gt} = \mu_{gt}\theta_g + (1 - \mu_{gt})(1 - \theta_g)$ is the predictive probability of winning. In addition μ_{gt} is the posterior belief at time t.

Learning rules are defined by the way in which μ_{gt} is updated relative to the 50-50 prior. Table 7 illustrates the posterior for each learning rule and each possible sequence of draws for $\theta = 0.7$.

Bayesian (BY)

Under Bayesian learning, individuals update their posterior beliefs (μ_t) using Bayes rule using all available observations from Option A at period t as follows:

$$
\begin{aligned}
\mu_{gt} &= Pr(\theta_g | k \ wins, t-k \ losses) \\
&= \frac{0.5(\theta_g)^k(1-\theta_g)^{t-k}}{0.5(\theta_g)^k(1-\theta_g)^{t-k} + 0.5(1-\theta_g)^k(\theta_g)^{t-k}}
\end{aligned}
$$

For Bayesian learners, the optimal choice for each decision would provide the best fit for their decisions.

First-1 (F1)

Under first-1 learning, individuals only take into consideration the observation in the first period to update their prior and ignores all subsequent observations. Thus individuals who exhibit this type of learning overweight initial information which persists into all their subsequent decisions. Updating is as follows:

$$
\begin{aligned}
\mu_g | Win &= Pr(\theta_g | Draw_1 \ is \ win) \\
&= \frac{0.5(\theta_g)^1(1-\theta_g)^0}{0.5(\theta_g)^1(1-\theta_g)^0 + 0.5(1-\theta_g)^1(\theta_g)^0} \\
&= \theta_g, \forall t > 0 \\
\mu_g | Loss &= Pr(\theta_g | Draw_1 \ is \ loss) \\
&= \frac{0.5(\theta_g)^0(1-\theta_g)^1}{0.5(\theta_g)^0(1-\theta_g)^1 + 0.5(1-\theta_g)^0(\theta_g)^1} \\
&= 1 - \theta_g, \forall t > 0
\end{aligned}
$$

Last-1 (L1)

Under last-1 learning, individuals only take into consideration the observation in previous period to update their prior and ignores all prior observations. Thus individuals who exhibit this type of learning overweight the most recent information. Updating is as follows:

$$\begin{aligned}
\mu_{g,t}|Win &= Pr(\theta_g|Draw_{t-1}\ is\ win) \\
&= \frac{0.5(\theta_g)^1(1-\theta_g)^0}{0.5(\theta_g)^1(1-\theta_g)^0 + 0.5(1-\theta_g)^1(\theta_g)^0}, \\
&= \theta_g \\
\mu_{g,t}|Loss &= Pr(\theta_g|Draw_{t-1}\ is\ loss) \\
&= \frac{0.5(\theta_g)^0(1-\theta_g)^{1-0}}{0.5(\theta_g)^0(1-\theta_g)^{1-0} + 0.5(1-\theta_g)^0(\theta_g)^{1-0}}. \\
&= 1-\theta_g
\end{aligned}$$

Random (RD)

In the random decision-making model there is no updating of the prior. Therefore $\mu_{gt} = 0.5\ \forall \theta, g, t$.

Status Quo (SQ)

In the status quo decision-making model, subjects make the same decision each time.

For each individual decision, we calculate x_{gt} under each learning rule, assuming a CRRA utility function with risk parameter computed from hypothetical risk decisions in the exit survey. For each individual, we run separate maximum likelihood estimation for each learning rule and evaluate the Bayesian Information Criterion, a goodness-of-fit measure, to evaluate which learning rule is the best fit for each individual: $BIC = -2LL + k * ln(n)$ where k is the number of parameters and n is the number of observations. We assign each individual to the learning rule model with the lowest BIC.

Table 8 shows the number of participants categorised into each learning rule. A small percentage (3.74%) of individuals fall into the status quo category. For 43.2% of participants, random decision-making best explains their choices. First-1 learning is second most likely, best explaining the decisions of 26.19% of individuals. Last-1 learning explains the decisions of about 21.43% of participants each. Only 5.44% percent of participants are classified as Bayesian learners. These findings are consistent with Gans, et al. (2007) and Barham et al. (2014), who both find a great deal of heterogeneity in learning rules and very little Bayesian learning.

We next explore what demographic characteristics may significantly predict learning styles using multinomial logit estimation. The results are shown in Table 9. Compared to random decision-making (the omitted category), individuals who are better educated (secondary or higher) are more likely to be Bayesian learners. Additionally individuals who own their farms and belong to a farm group are more likely to be classified as Last-1 learners. None of the other characteristics significantly predict learning categories.

5.2.4 Technology Adoption and Learning Rules

To analyze the effect of the different learning rules (SQ, BY, F1, L1 and RD) on adoption of of new farm inputs, we estimate the following regressions:

$$T_i = \alpha_0 + \sum_{LR=1}^{LR=5} \alpha_{LR} LR + \mathbf{X_i'}\gamma + \eta_i \tag{6}$$

where T_i is whether the farmer adopted a new farm input (crop, fertilizer, agrochemical, irrigation) in the last year, LR is learning rule: status-quo (1), bayesian (2), last-1 (3), first-1 (4) and random (5), $\mathbf{X_i'}$ is a vector of respondent characteristics (demographics, farming characteristics, risk and ambiguity proxies) and η_i is a stochastic error term. In the estimation, the omitted learning rule category is random decision-making. As all our dependent variables are Yes/No, we estimate the equation using a logit specification.

The results are shown in Table 10. We find that some learning styles are significantly related to adoption of new farm inputs. Relative to random decision-makers, indviduals who are are classified as Status Quo are less likely to have adopted any new farm input in the last year, however the result is only significant for crop adoption and more broadly any farm new input. This makes sense as these individuals are those who kept the same choice for all their decisions in the experiment, regardless of the draws from Option A, the value of λ or the value of θ suggesting that they are more like to stick with their choices for longer periods of time.

Our main result is the following. Last-1 learners, those who overweight information from the most recent draws, were significantly more likely to have adopted a new crop in the last year. These results suggest that depending on individual learning styles, choices of technology can be very different. Bayesian learning is positively correlated with new crop adoption and negatively correlated with fertilizer and agrochemical adoption, but none of these results are significant.[10]

6 Conclusions

In this paper we explored the connection between learning and technology choices in an artefactual field experiment conducted with farmers in Guyana. Participants made choices in one-armed bandit games between a known arm with fixed return and an unknown arm with uncertain return. To measure correctness of decision-making in the game, we constructed a number of indices, relative to proper Bayesian updating. We then used choices in the game to estimate learning styles which we correlate with farmer technology choices on their farms. We found heterogeneity in decision-making, including overweighting of initial and recent winning draws, and small amount of Bayesian learning.

[10]We also explored the effect of the different learning rules on stopping decisions of the farmers, whether they had ever stopped using a particular farm input, using similar estimation as above. The results (not shown) do not indcate any correlation between learning rules and stopping decisions.

Our main result suggesst that individuals who overweight recent draws are more likely to have adopted new farm inputs within the previous year. This behavior, similar to the well-known phenomonen of overweighting the likelihood, suggests that it is important to consider potential differences in how individuals learn when modeling technology adoption decisions. In this case, people could be making the right decision for the wrong reason.

References

Banks, J., Olson, M. & Porter, D. (1997). An experimental analysis of the bandit problem. *Economic Theory, 10*, 55-77.

Barham, B., Chavas, J., Fitz, D., Ríos-Salas, V. & Schechter, L. (2015). Risk, learning and technology adoption. *Agricultural Economics, 46*, 11-24.

Bergemann, D. & Hege, U. (1998). Dynamic venture capital financing, learning and moral hazard. *Journal of Banking and Finance, 22*, 703-735.

———— (2005). The financing of innovation: Learning and stopping. *RAND Journal of Economics, 36*(4), 719-752.

Binswanger, H. (1980). Attitudes toward risk: Experimental measurement in rural India. *American Journal of Agricultural Economics, 62*(3), 395-407.

Boyce, J., Bruner, D. & McKee, M. (2015). Strategic experimentation in the lab. *Managerian and Decision Economics* doi: 10.1002/mde.2723.

Bradt, R., Johnson, S., & Karlin, S. (1956). On sequential designs for maximizing the sum of n observations. *The Annals of Mathematical Statistics, 27*(4), 1060-1074.

Camerer, C. (1989). Does the basketball market believe in the hot hand?, *American Economic Review, 79*(5), 1257-1261.

Conley, T. G., & Udry, C. R. (2010). Learning about a new technology: Pineapple in Ghana. *American Economic Review, 100*(1), 35-69.

Copeland, A. (2007). Learning dynamics with private and public signals. *Economic Theory, 31*, 523-538.

Croson, R. & Sundali, J. (2005). The gambler's fallacy and the hot hand: empirical data from casinos. *Journal of Risk and Uncertainty, 30*(3), 195-209.

Engle-Warnick, J., Escobal, J., & Laszlo, S. (2009). How do additional alternatives affect individual choice under uncertainty?. *Canadian Journal of Economics, 42*, 113-140.

———— (2011). Ambiguity aversion and portfolio choice in small-scale Peruvian farming. *The B.E. Journal of Economic Analysis & Policy, 11*(1), 1-56.

Feder, G., Just, R. E., & Zilberman, D. (1985). Adoption of agricultural innovations in developing countries: A survey. *Economic Development and Cultural Change, 33*, 255-297.

Foster, A. & Rosenzweig, M. (1995). Learning by doing and learning from others: Human capital and technical change in agriculture. *Journal of Political Economy, 103*, 1176-1209.

———— (2010). Microeconomics of Technology Adoption. *Annual Review of Economics, 2*, 395-424.

Gans, N., Knox, G. & Croson, R. (2007) Simple models of discrete choice and their performance in bandit experiments. *Manufacturing & Service Operations Management, 9*, 383-408.

Grether, D. (1980). Bayes rule as a descriptive model: the representativeness heuristic. *Quarterly Journal of Economics, 95*(5), 37-557.

Harrison, G.W. & List, J.A. (2004). Field experiments. *Journal of Economic Literature, 42*, 1009-1055.

Jovanovic, B. (1979). Job search and the theory of turnover. *Journal of Political Economy, 87*, 972-990.

Jovanovic, B. & Nyarko, Y. (1996). Learning-by-doing and the choice of technology. *Econometrica, 64*6, 1299-1310.

Knight, J., Weir, S. & Woldehanna, T. (2003). The role of education in facilitating risk-taking and innovation in agriculture. *Journal of Development Studies*, *39*(6), 1-22.

Liu, E. M. (2013). Time to change what to sow: Risk preferences and technology adoption decisions of cotton farmers in China. *Review of Economics and Statistics*, *95*(4), 1386-1403.

Ross, N., Santos, P. & Capon, T. (2012). Risk, ambiguity and the adoption of new technologies: Experimental evidence from a developing economy. Conference, August 18-24, 2012, Foz do Iguacu, Brazil 126492, International Association of Agricultural Economists.

Rothchild, M. (1974). A two-armed bandit theory of market pricing. *Journal of Economic Theory*, *9*, 185-202.

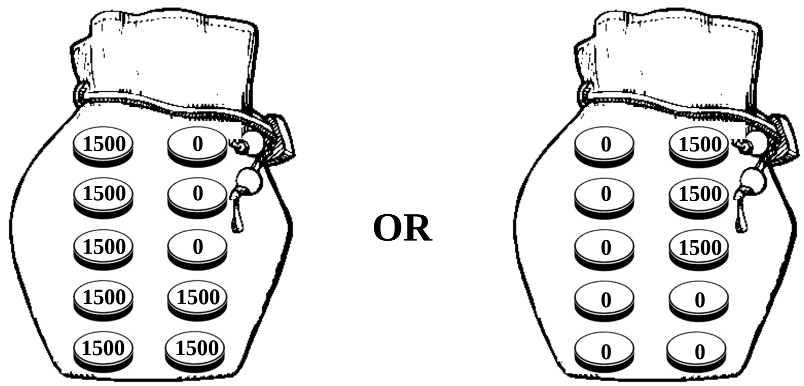
Rustichini, A. & Wolinsky, A. (1995). Learning about variable demand in the long run. *Journal of Dynamics and Control*, *19*, 1283-1292.

Figure 1: Instrument

OPTION A

ID:

50% chance **50% chance**



Draw Result 1: $________

Draw Result 2: $________

Draw Result 3: $________

Which bag do you think the chips came from?

☐ ☐

1

TASK 1: PLACE AN 'X' FOR YOUR CHOICE		
A	**OR**	**$250**
☐		☐

TASK 2: PLACE AN 'X' FOR YOUR CHOICE		
A	**OR**	**$375**
☐		☐

TASK 3: PLACE AN 'X' FOR YOUR CHOICE		
A	**OR**	**$600**
☐		☐

TASK 4: PLACE AN 'X' FOR YOUR CHOICE		
A	**OR**	**$250**
☐		☐

TASK 5: PLACE AN 'X' FOR YOUR CHOICE		
A	**OR**	**$375**
☐		☐

TASK 6: PLACE AN 'X' FOR YOUR CHOICE		
A	**OR**	**$600**
☐		☐

TASK 7: PLACE AN 'X' FOR YOUR CHOICE		
A	**OR**	**$250**
☐		☐

TASK 8: PLACE AN 'X' FOR YOUR CHOICE		
A	**OR**	**$375**
☐		☐

TASK 9: PLACE AN 'X' FOR YOUR CHOICE		
A	**OR**	**$600**
☐		☐

2

Figure 2: Number of Safe Choices

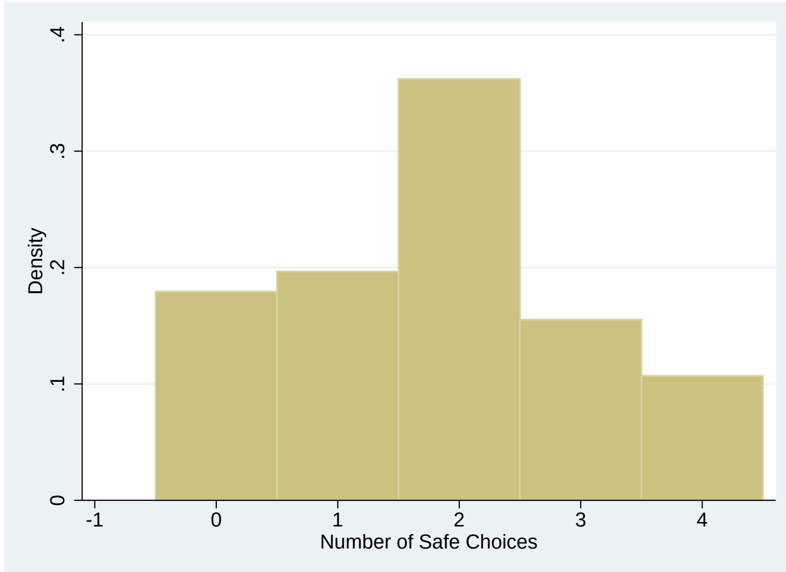


Figure 3: Number of Times Chose to Pay to Avoid Ambiguity

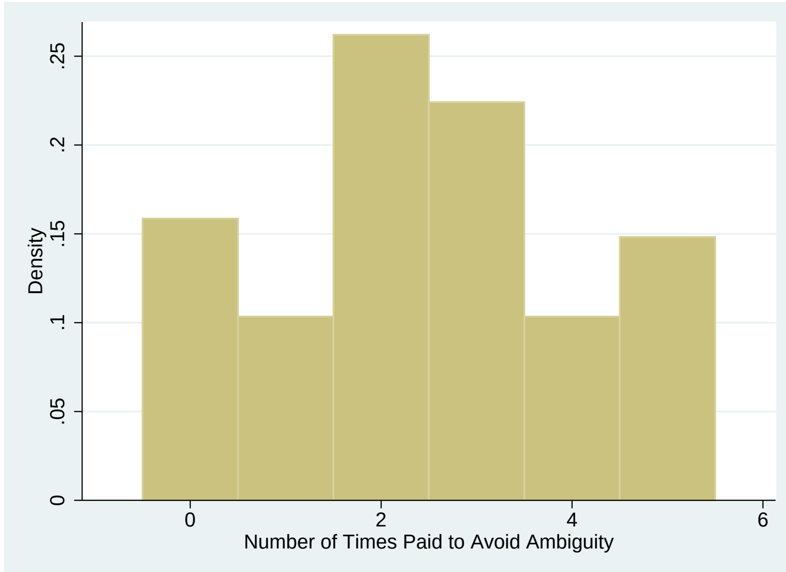


Figure 4: Game Level Index: $\theta = 0.7$

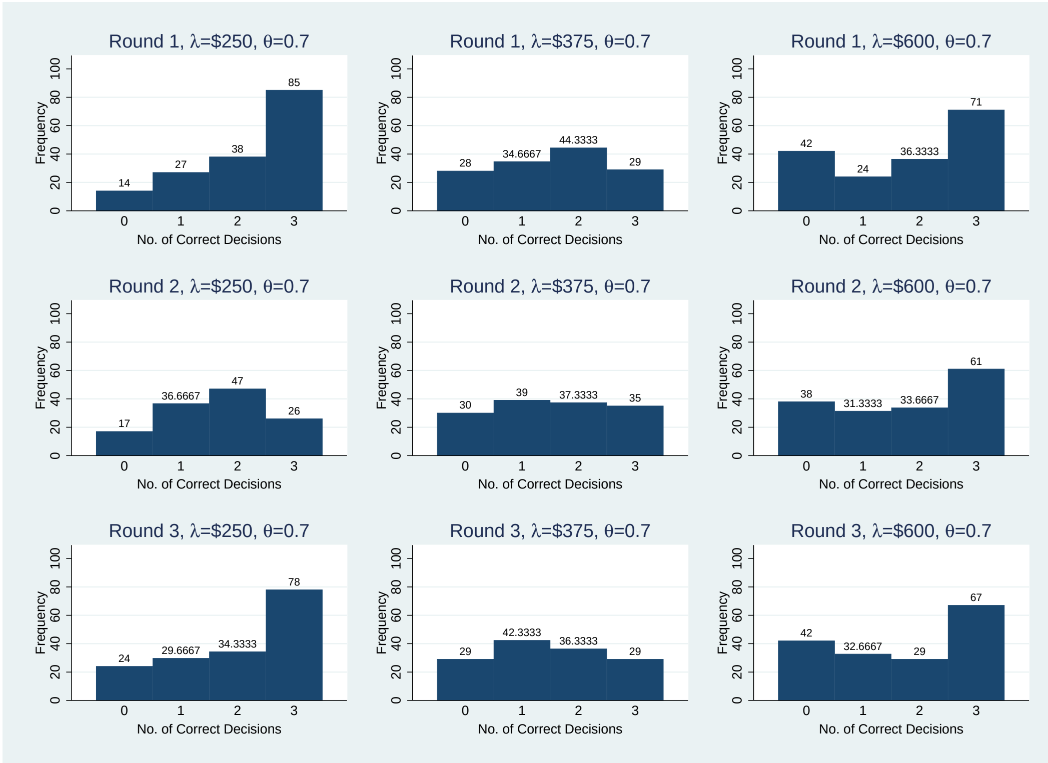


Figure 5: Game Level Index: $\theta = 0.9$

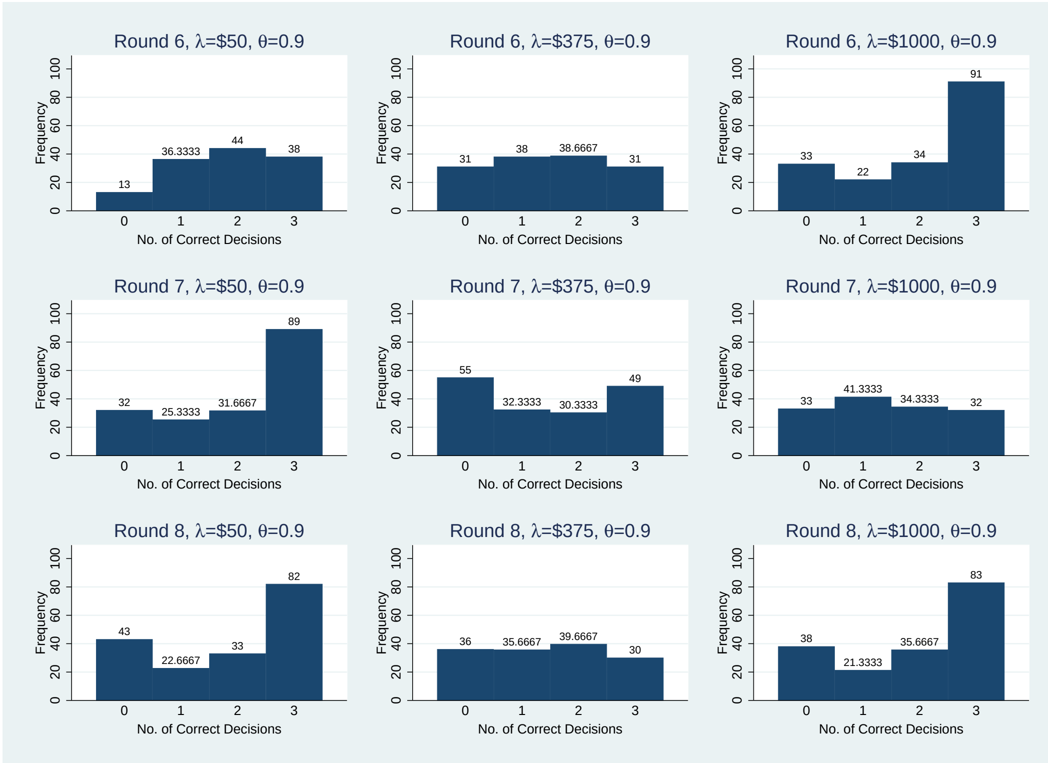


Figure 6: Game Level Index: $\theta = 1$ and $\theta = 0.5$

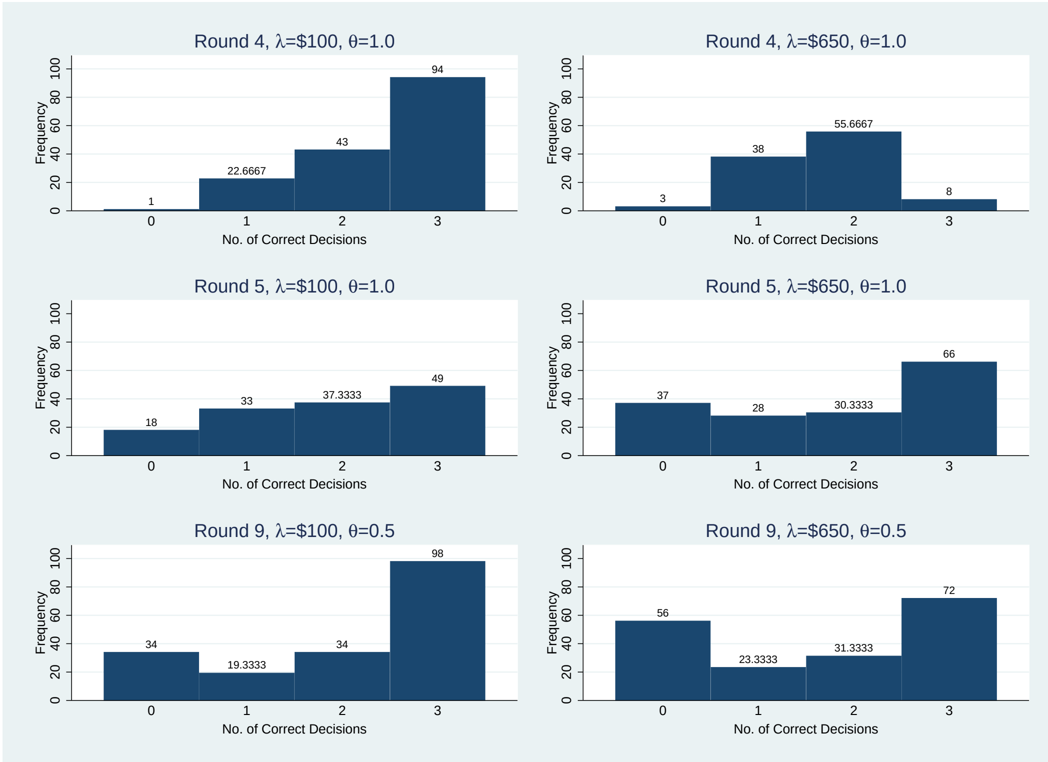


Figure 7: Individual Level Index

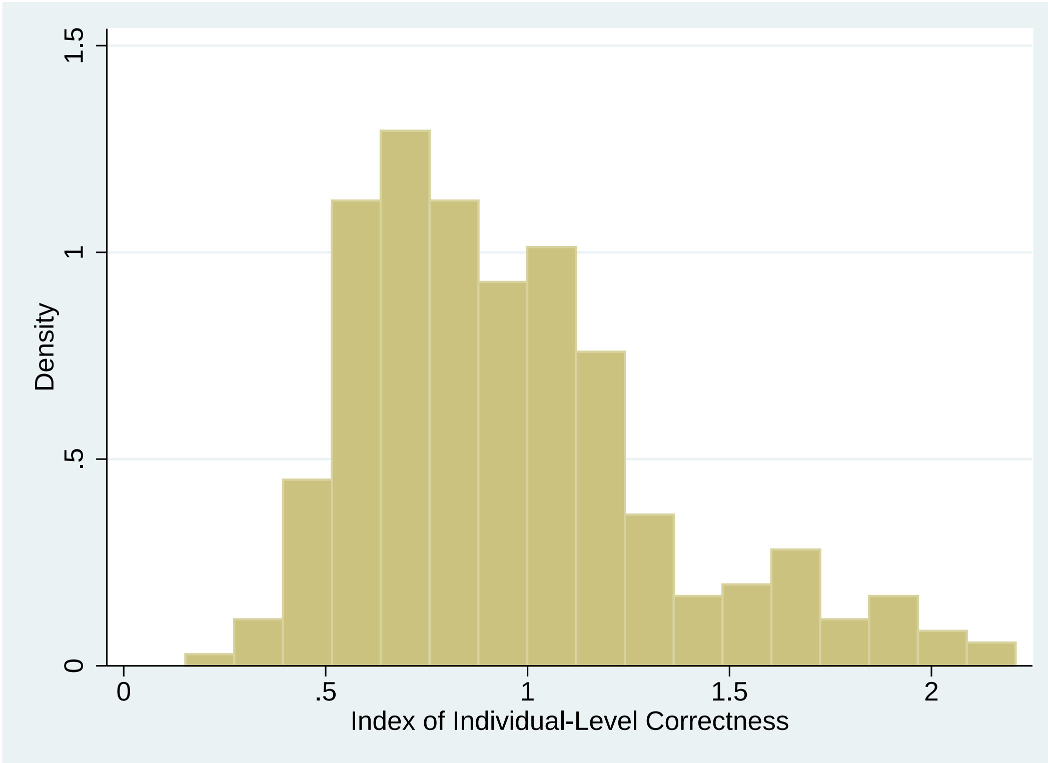


Table 1: Calibration of λ

θ	Strategy	Range for λ	Calibrated λ
1	KEP	>375	650
1	K1F	<375	100
0.9	KEP	≥ 669	1000
0.9	K1F	≥ 78 & <669	375
0.9	K2F	≥ 19 & <78	50
0.7	KEP	≥ 446	600
0.7	K1F	≥ 296 & <446	375
0.7	K2F	≥ 197 & <296	250
0.5	KEP	>375	650
0.5	AEP	<375	100

Table 2: Realizations of Outcomes, by Round

Round	θ	Realisation from Option A
1	0.7	1500, 0 1500
2	0.7	0, 0, 1500
3	0.7	0, 1500, 1500
4	1.0	1500, 1500, 1500
5	1.0	0, 0, 0
6	0.9	0, 0, 1500
7	0.9	1500, 1500, 1500
8	0.9	0, 1500, 1500
9	0.5	1500, 0, 1500

Table 3: Summary Statistics

Variable	Obs	Mean	Std. Dev
Participant Characteristics			
Age	290	43.84	14.35
Gender (Female=1)	290	0.31	0.47
Ethnicity			
East Indian	290	0.83	0.38
African	290	0.11	0.31
Other	290	0.06	0.24
Marital Status			
Married	290	0.80	0.40
Single	290	0.17	0.37
Widowed	290	0.03	0.18
Highest Education Level			
None/Nursery Schoool	290	0.02	0.14
Primary	290	0.45	0.50
Secondary	290	0.46	0.50
Post Secondary	290	0.06	0.24
Household Size	289	4.07	1.84
Flush Toilet (Yes = 1)	290	0.67	0.47
Drinking Water from Indoor Tap (Yes = 1)	289	0.37	0.48
Asset Index (0-1 Scale)	290	0.27	0.17
Farming Characteristics			
Number of Crops	281	3.58	1.72
Years Farming	288	20.46	13.78
Farm Size (Acres)	285	9.37	13.9
Land Ownership (Yes = 1)	287	0.74	0.44
Receives Technical Assistance (Yes = 1)	289	0.75	0.44
Member of Farm Group (Yes = 1)	290	0.21	0.41
Raises Livestock (Yes = 1)	290	0.61	0.49
Work Other than Farming (Yes = 1)	290	0.43	0.50
Access to Credit (1-5 Scale)			
Formal Sources	288	2.71	1.57
Informal Sources	288	2.63	1.36
Behavioral Measures			
Number of Safe Choices	290	1.81	1.21
Number of Times Chose to Pay to Avoid Ambiguity	290	2.46	1.58

Table 4: Top 10 crops cultivated

Crop Name	Frequency	Percent
Pepper	108	11.10
Bora (Yardlong Beans)	96	9.87
Okra	86	8.84
Rice	66	6.78
Boulanger (Eggplant)	61	6.27
Squash	52	5.34
Corilla (Bitter Melon)	51	5.24
Plantain	49	5.04
Tomato	46	4.73
Cassava	40	4.11

Table 5: Game Index, by θ and λ

		λ		
Trial	Theta	LOW	MEDIUM	HIGH
1	0.7	1.212	0.715	1.053
2	0.7	0.710	0.744	0.958
3	0.7	1.130	0.687	0.992
4	1.0	1.338	-	0.594
5	1.0	1.031	-	0.816
6	0.9	0.816	0.713	1.243
7	0.9	1.218	0.822	0.705
8	0.9	1.146	0.702	1.170
9	0.5	1.306	-	1.034

Table 6: Decision Level Indices

Index Option A when Should	Index λ when Should	Difference t-test	Wilcoxon Sum Rank
0.5914	0.5357	0.0557	2.570
(0.0123)	(0.0124)	(0.0217)**	[0.0105]**

Table 7: Posterior Distributions: Example $\theta = 0.7$

		Posterior (μ_t)		
Period	Draws	Bayesian	First-1	Last-1
0		0.5	0.5	0.5
1	0	0.3	0.3	0.3
2	0	0.16	0.3	0.3
0		0.5	0.5	0.5
1	0	0.3	0.3	0.3
2	1	0.5	0.3	0.7
0		0.5	0.5	0.5
1	1	0.7	0.7	0.7
2	0	0.5	0.7	0.3
0		0.5	0.5	0.5
1	1	0.7	0.7	0.7
2	1	0.84	0.7	0.7

Table 8: Learning Rules - Bayesian Information Criterion

Rule	Number	Percent
Bayesian	16	5.44
First-1	77	26.19
Last-1	63	21.43
Random	127	43.2
'Status Quo'	11	4.42

N=294

Table 9: Determinants of Learning Rules[&]

	Status Quo	Bayesian	Previous Draw	First Draw
Behavioral Measures				
Number of Safe Choices	0.65	0.88	0.92	1.02
Number of Times Chose to Pay to Avoid Ambiguity	1.05	0.818	1.07	1.17
Demographics				
Age	1.03	0.98	1.01	1.02
Sex (Female=1)	0.20	1.10	0.79	1.25
Secondary Education or Higher	0.73	4.17**	1.55	0.98
Asset Index	5.43	9.80	0.63	1.85
Farming Characteristics				
Land Ownership (Own=1) 1	0.90	0.88	2.48**	0.87
Farm Size (Acres)	0.94	0.99	0.99	1.01
No. of Years Farming	1.01	1.01	0.97*	0.99
Receives Technical Assistance	0.58	1.04	0.90	1.11
Belongs to Farm Group	1.10	1.43	2.32**	1.28
Wald χ^2 test				59.76*
Pseudo R-squared				0.0572

N=285. Multinomial logit relative risk ratios.

[&]Omitted learning rule is random decision-making.

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 10: New Technology Usage and Learning Rules

Variables	(1) New Crop	(2) New Fertilizer	(3) New Agrochemical	(4) Any New Technology
Learning Rules[&]				
Status Quo	-0.0406	-0.1315	-0.0226	-0.1954
	(0.050)	(0.059)**	(0.112)	(0.105)*
Bayesian	0.0191	-0.0850	-0.0332	-0.0485
	(0.149)	(0.077)	(0.063)	(0.229)
Previous Draw	0.1172	0.0144	-0.0091	-0.0018
	(0.050)**	(0.071)	(0.043)	(0.058)
First Draw	0.0823	-0.0256	-0.0433	-0.1051
	(0.053)	(0.055)	(0.055)	(0.076)
Behavioral Measures				
Number of Safe Choices	-0.0024	-0.0121	0.0088	0.0321
	(0.019)	(0.028)	(0.016)	(0.031)
Number of Times Chose to Pay to Avoid Ambiguity	-0.0179	-0.0027	-0.0370	-0.0199
	(0.017)	(0.016)	(0.019)**	(0.022)
Demographics				
Age	-0.0020	-0.0080	-0.0023	-0.0069
	(0.002)	(0.003)***	(0.003)	(0.004)*
Sex (Female=1)	-0.1016	-0.1525	-0.1100	-0.2132
	(0.052)*	(0.064)**	(0.059)*	(0.076)***
Secondary Education or More	-0.0410	0.0033	-0.0171	0.0086
	(0.031)	(0.048)	(0.050)	(0.048)
Asset Index	0.1754	0.3052	0.2807	0.3732
	(0.101)*	(0.108)***	(0.148)*	(0.198)*
Farming Characteristics				
Land Ownership (Own=1)	-0.0668	-0.0303	-0.0224	-0.0193
	(0.042)	(0.048)	(0.067)	(0.076)
Farm Size (Acres)	0.0028	0.0021	0.0036	0.0072
	(0.002)*	(0.001)	(0.002)**	(0.004)**
Years Farming	0.0042	0.0093	-0.0006	0.0082
	(0.002)**	(0.004)**	(0.003)	(0.004)**
Receives Technical Assistance	0.0333	-0.0186	-0.0530	-0.0583
	(0.044)	(0.055)	(0.051)	(0.068)
Belongs to Farm Group	0.0609	0.1160	0.0631	0.1711
	(0.064)	(0.061)*	(0.080)	(0.086)**
Observations	285	280	280	285
Wald χ^2 test	36.06***	39.28***	21.80	33.07***
Pseudo R-squared	0.1742	0.1486	0.0847	0.1260

[&]Omitted learning rule is random decision-making
Logit marginal effects with robust standard errors in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

1130, rue Sherbrooke Ouest, bureau 1400, Montréal (Québec) H3A 2M8
Tél. : 514-985-4000 • Téléc. : 514-985-4039
www.cirano.qc.ca • info@cirano.qc.ca

Centre interuniversitaire de recherche en analyse des organisations
Center for Interuniversity Research and Analysis on Organizations